

#9



TECH CENTER 1600/2900

AUG 13 2002

RECEIVED 1600

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/676,718A

DATE: 08/08/2002

TIME: 11:53:33

Input Set : A:\56113.app

Output Set: N:\CRF4\08082002\I676718A.raw

3 <110> APPLICANT: Gladyshev et al.

5 <120> TITLE OF INVENTION: MAMMALIAN SELENOPROTEIN DIFFERENTIALLY EXPRESSED IN TUMOR CELLS

7 <130> FILE REFERENCE: 4239-56113

9 <140> CURRENT APPLICATION NUMBER: US 09/676,718A

10 <141> CURRENT FILING DATE: 2000-09-28

12 <150> PRIOR APPLICATION NUMBER: PCT/US99/07560

13 <151> PRIOR FILING DATE: 1999-04-06

15 <150> PRIOR APPLICATION NUMBER: US 60/080,850

16 <151> PRIOR FILING DATE: 1998-04-06

18 <160> NUMBER OF SEQ ID NOS: 19

20 <170> SOFTWARE: PatentIn version 3.1

22 <210> SEQ ID NO: 1

23 <211> LENGTH: 162

24 <212> TYPE: PRT

25 <213> ORGANISM: Homo sapiens

27 <220> FEATURE:

28 <221> NAME/KEY: SITE

29 <222> LOCATION: (93)..(93)

30 <223> OTHER INFORMATION: Xaa is selenocysteine

33 <400> SEQUENCE: 1

35	Met	Ala	Ala	Gly	Pro	Ser	Gly	Cys	Leu	Val	Pro	Ala	Phe	Gly	Lys	Arg
36	1				5				10					15		
39	Leu	Leu	Leu	Ala	Thr	Val	Leu	Gln	Ala	Val	Ser	Ala	Phe	Gly	Ala	Glu
40				20				25						30		
43	Phe	Ser	Ser	Glu	Ala	Cys	Arg	Glu	Leu	Gly	Phe	Ser	Ser	Asn	Leu	Leu
44				35				40						45		
47	Cys	Ser	Ser	Cys	Asp	Leu	Leu	Gly	Gln	Phe	Asn	Leu	Leu	Gln	Leu	Asp
48				50				55						60		
51	Pro	Asp	Cys	Arg	Gly	Cys	Cys	Gln	Glu	Glu	Ala	Gln	Phe	Glu	Thr	Lys
52	65				70				75					80		
W--> 55	Lys	Leu	Tyr	Ala	Gly	Ala	Ile	Leu	Glu	Val	Cys	Gly	Xaa	Lys	Leu	Gly
56				85				90						95		
59	Arg	Phe	Pro	Gln	Val	Gln	Ala	Phe	Val	Arg	Ser	Asp	Lys	Pro	Lys	Leu
60				100				105						110		
63	Phe	Arg	Gly	Leu	Gln	Ile	Lys	Tyr	Val	Arg	Gly	Ser	Asp	Pro	Val	Leu
64				115				120						125		
67	Lys	Leu	Leu	Asp	Asp	Asn	Gly	Asn	Ile	Ala	Glu	Glu	Leu	Ser	Ile	Leu
68				130				135						140		
71	Lys	Trp	Asn	Thr	Asp	Ser	Val	Glu	Glu	Phe	Leu	Ser	Glu	Lys	Leu	Glu
72	145				150				155					160		
75	Arg	Ile														
79	<210>	SEQ ID NO:	2													
80	<211>	LENGTH:	1244													

## RAW SEQUENCE LISTING

DATE: 08/08/2002

PATENT APPLICATION: US/09/676,718A

TIME: 11:53:34

Input Set : A:\56113.app

Output Set: N:\CRF4\08082002\I676718A.raw

```

81 <212> TYPE: DNA
82 <213> ORGANISM: Homo sapiens
84 <220> FEATURE:
85 <221> NAME/KEY: CDS
86 <222> LOCATION: (5)..(493)
87 <223> OTHER INFORMATION:
90 <220> FEATURE:
91 <221> NAME/KEY: misc_feature
92 <222> LOCATION: (281)..(283)
93 <223> OTHER INFORMATION: TGA codon codes for selenocysteine
96 <400> SEQUENCE: 2
97 agcg atg gcg gct ggg ccg agt ggg tgt ctg gtg ccg gcg ttt ggg cta      49
98   Met Ala Ala Gly Pro Ser Gly Cys Leu Val Pro Ala Phe Gly Leu
99   1          5          10         15
101 cgg ttg ttg ttg gcg act gtg ctt caa gcg gtg tct gct ttt ggg gca      97
102 Arg Leu Leu Leu Ala Thr Val Leu Gln Ala Val Ser Ala Phe Gly Ala
103          20          25          30
105 gag ttt tca tcg gag gca tgc aga gag tta ggc ttt tct agc aac ttg      145
106 Glu Phe Ser Ser Glu Ala Cys Arg Glu Leu Gly Phe Ser Ser Asn Leu
107          35          40          45
109 ctt tgc agc tct tgt gat ctt ctc gga cag ttc aac ctg ctt cag ctg      193
110 Leu Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe Asn Leu Leu Gln Leu
111          50          55          60
113 gat cct gat tgc aga gga tgc tgt cag gag gaa gca caa ttt gaa acc      241
114 Asp Pro Asp Cys Arg Gly Cys Cys Gln Glu Glu Ala Gln Phe Glu Thr
115          65          70          75
117 aaa aag ctg tat gca gga gct att ctt gaa gtt tgt gga tga aaa ttg      289
118 Lys Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val Cys Gly Lys Leu
119 80          85          90
121 gga agg ttc cct caa gtc caa gct ttt gtt agg agt gat aaa ccc aaa      337
122 Gly Arg Phe Pro Gln Val Gln Ala Phe Val Arg Ser Asp Lys Pro Lys
123 95          100         105         110
125 ctg ttc aga gga ctg caa atc aag tat gtc cgt ggt tca gac cct gta      385
126 Leu Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg Gly Ser Asp Pro Val
127          115         120         125
129 tta aag ctt ttg gac gac aat ggg aac att gct gaa gaa ctg agc att      433
130 Leu Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala Glu Glu Leu Ser Ile
131          130         135         140
133 ctc aaa tgg aac aca gac agt gta gaa gaa ttc ctg agt gaa aag ttg      481
134 Leu Lys Trp Asn Thr Asp Ser Val Glu Glu Phe Leu Ser Glu Lys Leu
135          145         150         155
137 gaa cgc ata taa atcttgctta aattttgtcc tatccttttg ttacottatc      533
138 Glu Arg Ile
139          160
141 aaatgaaata ttacagcacc tagaaaataa tttagttttg cttgottcca ttgatcagtc      593
143 ttttacttga ggcattaaat atctaattaa atcgtgaaat ggcagtatag tccatgatat      653
145 ctaaggagtt ggcaagctta acaaaaccca ttttttataa atgtccatcc tctgcattt      713
147 gttgatacca ctaacaaaat gctttgtaac agacttgcgg ttaattatgc aaatgatagt      773
149 ttgtgataat tgggtccagtt ttacgaacaa cagatttcta aattagagag gttaacaaga      833

```

## RAW SEQUENCE LISTING

DATE: 08/08/2002

PATENT APPLICATION: US/09/676,718A

TIME: 11:53:34

Input Set : A:\56113.app

Output Set: N:\CRF4\08082002\I676718A.raw

```

151 cagatgatta ctatgcctca tgtgctgtgt gctctttgaa aggaatgaca gcagactaca      893
153 aagcaaataa gatatactga gcctcaacag attgcctgct cctcagagtc tctcctatct      953
155 ttgtattacc cagctttctt ttttaatacaa atgttattta tagtttacaa tgaatgcact    1013
157 gcataaaaac tttgtagctt cattattgta aaacatattc aagatcctac agtaagagtg    1073
159 aaacattcac aaagatttgc gttaatgaag actacacaga aaacctttct agggatttgt    1133
161 gtggatcaga tacatacttg gcaaattttt gagttttaca ttcttacaga aaagtccatt    1193
163 taaaagtgat catttgtaag accaaaatat aaataaaaag tttcaaaaat c              1244
166 <210> SEQ ID NO: 3
167 <211> LENGTH: 489
168 <212> TYPE: DNA
169 <213> ORGANISM: Homo sapiens
171 <220> FEATURE:
172 <221> NAME/KEY: CDS
173 <222> LOCATION: (1)..(489)
174 <223> OTHER INFORMATION:
177 <220> FEATURE:
178 <221> NAME/KEY: misc_feature
179 <222> LOCATION: (277)..(279)
180 <223> OTHER INFORMATION: TGA codon codes for selenocysteine
183 <400> SEQUENCE: 3
184 atg gcg gct ggg ccg agt ggg tgt ctg gtg ccg gcg ttt ggg cta cgg      48
185 Met Ala Ala Gly Pro Ser Gly Cys Leu Val Pro Ala Phe Gly Leu Arg
186 1          5          10          15
188 ttg ttg ttg gcg act gtg ctt caa gcg gtg tct gct ttt ggg gca gag      96
189 Leu Leu Leu Ala Thr Val Leu Gln Ala Val Ser Ala Phe Gly Ala Glu
190          20          25          30
192 ttt tca tcg gag gca tgc aga gag tta ggc ttt tct agc aac ttg ctt      144
193 Phe Ser Ser Glu Ala Cys Arg Glu Leu Gly Phe Ser Ser Asn Leu Leu
194          35          40          45
196 tgc agc tct tgt gat ctt ctc gga cag ttc aac ctg ctt cag ctg gat      192
197 Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe Asn Leu Leu Gln Leu Asp
198          50          55          60
200 cct gat tgc aga gga tgc tgt cag gag gaa gca caa ttt gaa acc aaa      240
201 Pro Asp Cys Arg Gly Cys Cys Gln Glu Glu Ala Gln Phe Glu Thr Lys
202 65          70          75          80
204 aag ctg tat gca gga gct att ctt gaa gtt tgt gga tga aaa ttg gga      288
205 Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val Cys Gly          Lys Leu Gly
206          85          90          95
208 agg ttc cct caa gtc caa gct ttt gtt agg agt gat aaa ccc aaa ctg      336
209 Arg Phe Pro Gln Val Gln Ala Phe Val Arg Ser Asp Lys Pro Lys Leu
210          100          105          110
212 ttc aga gga ctg caa atc aag tat gtc cgt ggt tca gac cct gta tta      384
213 Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg Gly Ser Asp Pro Val Leu
214          115          120          125
216 aag ctt ttg gac gac aat ggg aac att gct gaa gaa ctg agc att ctc      432
217 Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala Glu Glu Leu Ser Ile Leu
218          130          135          140
220 aaa tgg aac aca gac agt gta gaa gaa ttc ctg agt gaa aag ttg gaa      480
221 Lys Trp Asn Thr Asp Ser Val Glu Glu Phe Leu Ser Glu Lys Leu Glu

```

## RAW SEQUENCE LISTING

DATE: 08/08/2002

PATENT APPLICATION: US/09/676,718A

TIME: 11:53:34

Input Set : A:\56113.app

Output Set: N:\CRF4\08082002\I676718A.raw

```

222      145      150      155
224 cgc ata taa
225 Arg Ile
226 160
229 <210> SEQ ID NO: 4
230 <211> LENGTH: 136
231 <212> TYPE: PRT
232 <213> ORGANISM: Homo sapiens
234 <220> FEATURE:
235 <221> NAME/KEY: SITE
236 <222> LOCATION: (67)..(67)
237 <223> OTHER INFORMATION: Xaa is selenocysteine
240 <400> SEQUENCE: 4
242 Ser Ala Phe Gly Ala Glu Phe Ser Ser Glu Ala Cys Arg Glu Leu Gly
243 1      5      10      15
246 Phe Ser Ser Asn Leu Leu Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe
247      20      25      30
250 Asn Leu Leu Gln Leu Asp Pro Asp Cys Arg Gly Cys Cys Gln Glu Glu
251      35      40      45
254 Ala Gln Phe Glu Thr Lys Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val
255      50      55      60
W--> 258 Cys Gly Xaa Lys Leu Gly Arg Phe Pro Gln Val Gln Ala Phe Val Arg
259 65      70      75      80
262 Ser Asp Lys Pro Lys Leu Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg
263      85      90      95
266 Gly Ser Asp Pro Val Leu Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala
267      100     105     110
270 Glu Glu Leu Ser Ile Leu Lys Trp Asn Thr Asp Ser Val Glu Glu Phe
271      115     120     125
274 Leu Ser Glu Lys Leu Glu Arg Ile
275      130     135
278 <210> SEQ ID NO: 5
279 <211> LENGTH: 21
280 <212> TYPE: DNA
281 <213> ORGANISM: Artificial Sequence
283 <220> FEATURE:
284 <223> OTHER INFORMATION: Primer
286 <400> SEQUENCE: 5
287 atggcggctg ggccgagtg g
289      21
290 <210> SEQ ID NO: 6
291 <211> LENGTH: 21
292 <212> TYPE: DNA
293 <213> ORGANISM: Artificial Sequence
295 <220> FEATURE:
296 <223> OTHER INFORMATION: Primer
298 <400> SEQUENCE: 6
299 taatatgcgt tccaactttt c
301      21
302 <210> SEQ ID NO: 7
303 <211> LENGTH: 21

```

## RAW SEQUENCE LISTING

DATE: 08/08/2002

PATENT APPLICATION: US/09/676,718A

TIME: 11:53:34

Input Set : A:\56113.app

Output Set: N:\CRF4\08082002\I676718A.raw

```

304 <212> TYPE: DNA
305 <213> ORGANISM: Artificial Sequence
307 <220> FEATURE:
308 <223> OTHER INFORMATION: Primer
310 <400> SEQUENCE: 7
311 tctgcttttg gggcagagtt t                                     21
314 <210> SEQ ID NO: 8
315 <211> LENGTH: 1216
316 <212> TYPE: DNA
317 <213> ORGANISM: Mus musculus
319 <220> FEATURE:
320 <221> NAME/KEY: CDS
321 <222> LOCATION: (11)..(490)
322 <223> OTHER INFORMATION:
325 <220> FEATURE:
326 <221> NAME/KEY: misc_feature
327 <222> LOCATION: (287)..(289)
328 <223> OTHER INFORMATION: TGA codon codes for selenocysteine
331 <400> SEQUENCE: 8
332 gaccgcaggg atg gcg gca ggg cag ggt ggg tgg ctg cgg cca gct ctg       49
333      Met Ala Ala Gly Gln Gly Gly Trp Leu Arg Pro Ala Leu
334      1          5          10
336 ggg ctg cgc ttg ctg ctg gcg act gcg ttt caa gcg gtg tct gct ctg       97
337 Gly Leu Arg Leu Leu Leu Ala Thr Ala Phe Gln Ala Val Ser Ala Leu
338      15          20          25
340 ggg gca gag ttt gcg tca gag gca tgc aga gag ttg ggt ttc tcc agc      145
341 Gly Ala Glu Phe Ala Ser Glu Ala Cys Arg Glu Leu Gly Phe Ser Ser
342 30          35          40          45
344 aac ttg ctc tgc agc tct tgc gat ctt ctt gga cag ttt aat ctg ctc      193
345 Asn Leu Leu Cys Ser Ser Cys Asp Leu Leu Gly Gln Phe Asn Leu Leu
346      50          55          60
348 cca ctg gac cct gtt tgc aga ggg tgc tgt cag gaa gaa gca caa ttt      241
349 Pro Leu Asp Pro Val Cys Arg Gly Cys Cys Gln Glu Glu Ala Gln Phe
350      65          70          75
352 gaa acc aaa aag ctg tat gca gga gcc atc ctt gaa gtc tgc gga tga      289
353 Glu Thr Lys Lys Leu Tyr Ala Gly Ala Ile Leu Glu Val Cys Gly
354      80          85          90
356 aaa ttg ggg agg ttc cct caa gtc caa gct ttt gtc aga agt gat aaa      337
357 Lys Leu Gly Arg Phe Pro Gln Val Gln Ala Phe Val Arg Ser Asp Lys
358      95          100          105
360 ccc aaa ctc ttc aga ggt cta cag atc aag tat gtt cga ggc tca gac      385
361 Pro Lys Leu Phe Arg Gly Leu Gln Ile Lys Tyr Val Arg Gly Ser Asp
362      110          115          120
364 cct gta cta aag ctt ttg gac gac aac ggg aac att gct gaa gaa cta      433
365 Pro Val Leu Lys Leu Leu Asp Asp Asn Gly Asn Ile Ala Glu Glu Leu
366 125          130          135          140
368 agc atc ctc aaa tgg aac aca gac agt gtg gaa gag ttc ctg agc gag      481
369 Ser Ile Leu Lys Trp Asn Thr Asp Ser Val Glu Glu Phe Leu Ser Glu
370      145          150          155

```

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 08/08/2002  
PATENT APPLICATION: US/09/676,718A      TIME: 11:53:35

Input Set : A:\56113.app  
Output Set: N:\CRF4\08082002\I676718A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 93  
Seq#:4; Xaa Pos. 67  
Seq#:9; Xaa Pos. 93  
Seq#:16; Xaa Pos. 129

## VERIFICATION SUMMARY

DATE: 08/08/2002

PATENT APPLICATION: US/09/676,718A

TIME: 11:53:35

Input Set : A:\56113.app

Output Set: N:\CRF4\08082002\I676718A.raw

L:55 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:80  
L:258 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:64  
L:434 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:80  
L:575 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:128